Risk Analysis in Environmental and Occupational Health
September 1, 1987

Uncertainties in Predicting Human Risks

Edmund Crouch

1. Background Information

It is useful to bear in mind a few sobering facts about total populations at risk, and the normal total risk of death and of dying of cancer. For the U.S., the total population is about 240 million, while the annual number of deaths is about 2 million per year and the annual number of cancer deaths is about 400 thousand. These figures imply an annual average total risk of death of about 10^{-2} (1 percent per year), and a lifetime risk of cancer of about 0.2 (20 percent, or $200,000 \times 10^{-6}$), estimates you can obtain simply by dividing one figure by another.

Of course, simply dividing one by another is not a particularly accurate way of computing such estimates -- one should do the correct thing and take the age structure of the population into account, and the variation of risks with age, and so on. But even when you do precisely that, the average lifetime risk of cancer comes out to be about 20 to 25 percent. We can expect this figure to get higher as the expectation of life increases, and as other causes of death are eliminated (assuming -- pessimistically --that most cancers cannot be eliminated). It is mainly the increase in expectation of life which has made cancer such a prominent cause of death in the (historically) recent past, because cancers tend to be diseases of old age.

For many cancers it is found that the death rate varies as a power of age:-

$$\text{rate} \sim \text{age}^n$$

where the exponent n is in the range 4 to 11. For such cancers, this pattern seems to hold over the age range from about 30 to 65. At lower ages the rates tend to be very small but almost independent of age (and the cancers may be completely different diseases in youngsters), while at higher ages the reported death rates are lower than would be predicted by this sort of formula - and in some cases the reported death rates are actually lower for old enough groups. It is unclear whether these reductions in death rates in the elderly are real, or are simply due to a difference in the accuracy of diagnosis and reporting. It is also possible that the reduction in reported death rates is real, but is due to the winnowing out of the population of those who are susceptible to these particular cancers, leaving a core of more resistant individuals.

The major exceptions to the power law variation of death rate with age are the cancers which are known to be hormonally dependent (e.g. breast cancer), or are highly curable (skin cancers), or in which the natural progression is altered by intervention (e.g. a high proportion of women have had hysterectomies by age 65, so that they cannot be at

2025545851

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

risk of uterine cancers thereafter).

With this age variation of risk of cancer understood, we can now oversimplify again and quote a lifetime average annual risk for cancer, obtained simply by dividing the lifetime risk by an average lifetime of about 70 years. This give an average annual risk of about $2 - 3 \times 10^{-3}$. Notice that we are here averaging over a lifetime -- the figure is not meant to imply that the risk is the same in each year of life -- we have just seen that it varies drastically with age.

When discussing the risks of carcinogens, the same caveats have to be borne in mind. We usually attempt to estimate a lifetime risk, but may express this, for comparison purposes, as an annual average risk. For an individual exposed continuously to a carcinogen, we would expect that the risk of cancer increases with age in a fashion similar to the risk of other (naturally occurring) cancers.

There is another reason also for quoting an annual average risk obtained by averaging over a lifetime. When estimating risks of carcinogens, one is often interested in the response of a population to exposure to the carcinogen. In this case, one should strictly (if it were possible) estimate what the effects at all future times would be on individuals of different ages at the times of exposure. The effects at all future times on the whole population would then be an average over the effects on all the individuals in the population (who were of different ages at the times of exposure.

Thus, to obtain an estimate of the effects on a population, one implicitly performs an average over the age groups present in the population. If the population were stationary (and if certain other conditions were fulfilled) this average would be the same as an average over a lifetime. This explains the usefulness of a lifetime average, since one may argue that the differences between population and lifetime averages are small compared with other uncertainties inherent in all the procedures we will describe later.

The preceding discussion must be considered only a heuristic argument for accepting a lifetime average as being useful. In practice, people will be exposed at different ages, and for varying periods, to different amounts of carcinogens. All these differences (and many more besides) will affect the probability of carcinogenesis for each of them.

2. Known Human Carcinogens

There is now good evidence that human exposure to certain materials can, under certain conditions, increase the rate of human cancer. The evidence comes from various types of epidemiological investigation (discussed in other talks in this course). In all cases, exposures to these materials has been high, compared with population exposures, and the population exposed has been small compared with the total U.S. population. The resultant risks to those exposed has been substantial.

The following table indicates a few of these materials, and the types of cancer which have been caused in humans by exposure to them.

2025545852

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

Also shown are the "natural" rates for such cancers, expressed in terms of lifetime risk and annual average risk.

Material/Action	Site or Type	Lifetime Risk (In absence of exposures)	Annual Average Risk
4-Aminobiphenyl Auramine manufacture Benzidine Chlornaphazine Cyclophosphamide 2-Naphthylamine	Bladder	5×10^{-3}	7×10^{-5}
Arsenic (compounds) Asbestos BCME CCME Chromium (VI compounds) Mustard gas Nickel refining	Lung	4×10^{-2} (Pop^{n}. ave.)	6×10^{-4}
Arsenic PUVA Soots, Tars, Mineral oils	Skin	3×10^{-3} (Deaths!)	4×10^{-5}
Vinyl chloride	Liver	1×10^{-3}	2×10^{-5}
DES (In Utero)	Vagina	7×10^{-3}	9×10^{-5}
Benzene Myleran Chlorambucil Melphalan	Leukemia	8×10^{-3}	1×10^{-4}

Typically, in epidemiological studies, a relative risk of >2 is required in order to detect any effect. Thus the (epidemiologically) discoverable population average human risks are $> 10^{-5}$ per year, or 10^{-3} per lifetime, and probably much larger. For the small subgroups of the population usually available for study, the observable risks are generally much larger.

3. Target Risks. The Necessity of Extrapolation.

When considering the size of acceptable risks to the public at large, the usual targets are much smaller than the discoverable risks discussed above. Typically they will be of order $< 10^{-6}$ per year. Note that the EPA and the FDA set targets of order 10^{-6} per lifetime, that is, of order 10^{-8} per year.

It must also be borne in mind that there are a large number of materials which are of potential interest. The Chemical Abstracts

2025545853

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

Service (CAS) has now given names to well over six million distinct chemicals which have been mentioned in scientific literature, and there have been various estimates of the number (around 50,000) of chemicals in general commercial use.

With such numbers, it should be immediately apparent that there are just too many time, money and logistical constraints to directly detecting any adverse effects from such a plethora of materials to which humans may be exposed. Notice that a risk of 10^{-6} per lifetime corresponds to a rate of about 3 per year in the whole U.S. population. Thus, even if the whole U.S. population were exposed to some material causing a risk of death of 10^{-6} per lifetime, the resulting deaths would be statistically indistinguishable in the usual two million deaths per year (unless there were something extremely unusual about the deaths).

Extrapolation is therefore essential in order to estimate the sizes of risks, and hence be in a position to demand that risks be reduced to the levels mentioned. The fundamental observation on which such extrapolation is based is that:

HUMAN CARCINOGEN -->implies--> ANIMAL CARCINOGEN

In other words, every known material which has been shown to be a human carcinogen is also known to cause tumors in animals under suitable conditions. The only current possible exception to this is arsenic, but it is quite plausible that this is simply because it has not been tested adequately.

This observation is not very useful in itself, but what is done in order to allow risk assessments is to assume its converse:

ANIMAL CARCINOGEN -->implies--> HUMAN CARCINOGEN

and to work from here. This assumption is not unreasonable, in view of what is known about carcinogenesis - although it is something which can be argued about in specific cases.

4. The Nature of Carcinogenesis.

In what follows, it is useful to keep in mind some information about the process of carcinogenesis. This information has been derived from studies of humans and animals, and from experiments performed in vivo or in vitro. It is based partly on experimental studies, and partly on theoretical ideas suggested by those studies.

(a.) Cancers arise from one (or more) individual cell(s) which have gone "out of control" in some way - the cell becomes immortal, with no limit on the number of cell divisions, and the usual constraints on cell division no longer apply. A cell may pass through several stages before reaching this state.

(b.) The underlying cause of such behavior is probably some effect(s) on the genetic material of the cell, but the exact mechanism(s) is (are) unknown.

2025545854

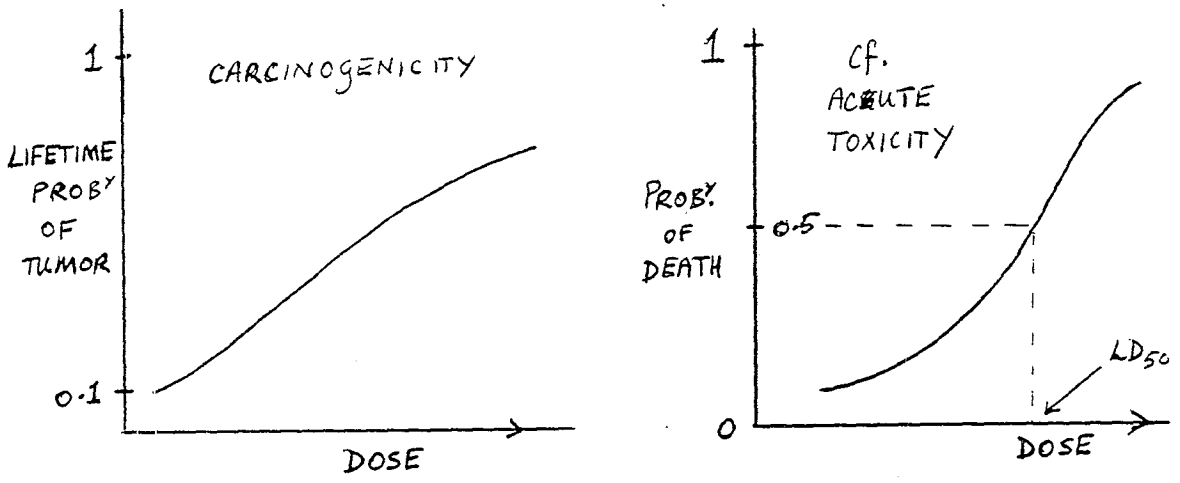
Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

(c.) The occurrence of such events appears to be a random process at some level. One cannot tell which individual cell or animal or person will be affected. Hence we talk about the PROBABILITIES of cancer - the chance that some event will occur.

(d.) When we feed materials to experimental animals, these probabilities depend on various factors which can be manipulated. For example, they vary with:

The total AMOUNT of material (the total dose)
The AGE at which dosing takes place
The RATE OF APPLICATION, or the time over which dosing continues
OTHER FACTORS (some known -- stress, dietary factors, ...,
others unknown)

We therefore expect, and in practice observe, DOSE-RESPONSE curves. Such dose-response curves are fundamental in extrapolating risks to humans. I like to draw an analogy to the similar problem of extrapolation which arises for acute toxicity -- in both cases, we have measurement difficulties at low doses, and in both cases there is some sort of dose-response relationship (which I deliberately leave vague for now):



(e.) Evidently there will be some AGE STRUCTURE to the probabilities of cancer. As mentioned, for many cancers in humans the death rate from cancers increases with a power of age. In experimental studies involving long term feeding of rodents, the same sort of age structure is found for the incidence of tumors. A "LIFETIME" probability thus depends on how you measure it - the usual practice is to assume a "standard" lifetime of ~70 years for humans and ~2 years for rodents.

(f.) At high enough doses (i.e.. at high RESPONSES one sees interactions between different materials in both animal experiments and in human data (e.g. smoking and alcohol consumption, smoking and radon exposure, smoking and asbestos exposure). The effect of such interactions is to make the effect of two or more materials different from the sum of the effects of the materials individually (at the same

2025545855

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

doses).

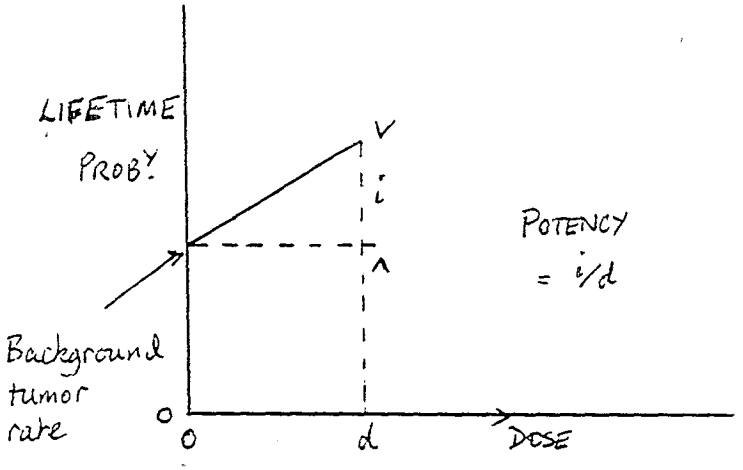
(g.) It is not possible to make direct measurements of what happens at low doses (i.e. at LOW RESPONSES). In this context, low dose means a dose at which the response probability is < 0.1 usually, and < 0.01 certainly. Any attempt at studying lower doses runs up against problems of logistics, cost and the background cancer rate.

(h.) The shape of dose-response curves assumed for the low dose regions are thus based on:

Theoretical ideas
Prejudice
Guesswork

For performing risk assessments for human safety purposes, there is naturally a prejudice to be conservative.

It is generally agreed that assuming LINEARITY between dose and response (for our discussion, this means the lifetime probability of a cancer) at low enough doses is CONSERVATIVE. This assumption is made in a theoretical way -- it is assumed that the true relationship between dose and response lies, at low enough doses, entirely below (or at worst on) a linear curve joining the response at zero dose (background) with the response at some higher (but still low) dose.



Typically, the background rate is of order 10^{-4} to 10^{-1}, and we are interested in excesses over the background of order 10^{-6} to 10^{-4}, so this diagram is not to scale. It is useful to define the POTENCY of a carcinogen as the ratio of excess lifetime probability of cancer to the dose causing that excess (at low enough doses). On the diagram, this is the ratio i/d. The potency is thus the slope of the dose-response curve at low enough dose, and we have the basic equation:

EXCESS RISK = POTENCY x DOSE

There is reasonable evidence that some mechanisms of carcinogenesis result in a THRESHOLD -- i.e. that there is some (threshold) dose below which the excess incidence of cancer is much lower than would be

2025545856

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

predicted by a linear extrapolation from doses above the threshold, and possibly that the excess incidence of cancer is literally zero below such a threshold (excess, here, means excess over the background occurrence of cancer). Some of the evidence for such mechanisms comes from observation of the dose-response curves in experimental situations -- the experiments on saccharin provide a good example. However, there is still the possibility that a linear mechanism may still operate at low enough doses, and so any human risk assessment has to take that possibility into account.

5. The Standard Animal Test.

The requirements for a "standard" animal test are quite severe. The animals involved have to be as similar to humans as possible -- in metabolism, in being omnivorous, in their sensitivity to chemicals, for example -- yet as different as possible in their life span and cost of upkeep (so that we can get results in a reasonable time at a reasonable cost). In practice, there is little option but to use standard laboratory animals. The usual choices are rodents -- rats and mice; with occasional tests being performed on golden hamsters or guinea pigs. Other animals (e.g. gerbils) have been proposed, but for now the experience built up in handling laboratory rodents is a strong incentive for continuing their use despite certain known disadvantages. Any change would now have to be done gradually, and with much cross checking with previous results.

It is now standard to require tests to be performed in at least two species (practically always rats and mice) and on both sexes, in case one or the other species or sex is peculiarly resistant to the material under test. A compromise has to be made over the number of animals to test. It would be desirable to have as many as logistically possible, to increase the statistical sensitivity of the experiment; but as few as possible to minimize the costs of testing (since there is always another material to test). The current recommendation is for at least 50 per group of similarly treated animals.

There is a similar trade-off between costs and the number of dose levels to test in a given experiment. The current recommendation is to have at least three dose groups -- an undosed group (the control group), a group tested at the maximum tolerated dose (MTD) of the material under test, and the third group tested at some intermediate dose (usually 1/4 to 1/2 of the MTD).

The MTD of a material is roughly defined to be as much as possible, but not enough to kill off the animals early or to cause too large other overt effects (like loss of weight). The reason for using it in these experiments is to increase the sensitivity, on the basis that giving more of something is more likely to produce a response if any response if going to happen at all. The sensitivity has to be as high as possible, since the observable responses are of the order 10^{-1} (10%) while the risks of interest are of order 10^{-6} (100,000 times smaller). The alternative way of increasing sensitivity is to increase the number of animals tested (within reason), but this only increases sensitivity in proportion to the square root of the numbers tested, while increasing

2025545857

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

the dose gives an increase in sensitivity roughly proportional to the dose. Clearly the latter is most cost effective.

Even with such a minimum design, there are:

3 dose groups x 2 sexes x 2 species x 50 animals per group

giving a minimum of 600 animals per experiment. All the animals have to be carefully housed (under standard conditions), cared for, and individually tracked throughout their two year lifetime. They are then sacrificed and a large number of their tissues examined individually. None of this comes cheap -- the cost of such an experiment is unlikely to be less than $200,000, and may run above $1,000,000.

It should be noted that the type of experiment detailed here is the minimum considered necessary to answer a YES/NO question: Is this material carcinogenic under the conditions of this standard bioassay? The experimental design and analyses performed are designed to be unlikely to answer YES if there is no carcinogenic action present (so that the experiments have low alpha error), but they can easily answer NO even in the presence of carcinogenic action. This sort of test is exactly what is required, of course, if one is interested in identifying materials which are surely carcinogens; in order to study their mechanism of action for example -- one doesn't want to accidentally end up with a material with no carcinogenic action.

I would submit, however, that for the purposes of protection of public health, the questions asked of the tests are entirely the wrong way round. For protecting public health, one should surely ask not whether this material is almost surely a carcinogen, but how strong a carcinogen it could be, given the results of the experiment. The fact that the same sort of analysis is applied now as in the past is perhaps a combination of accident and inertia, but one has to admit that, for the most part, the methodology has been largely successful so far.

6. Raw Results - and what to do with them.

Having spent 2 years performing the experiment described above, what output do we get? When the animals are sacrificed, they are dissected and a whole list of tissues examined, both macroscopically and microscopically. All lesions, whether related to cancer or not, are noted down and usually (nowadays) recorded in some sort of computer database. The pathologists performing the examinations usually use some sort of standardized nomenclature for what they observe -- for example, the National Toxicology Program uses a modified version of the Systematized Nomenclature for Pathology (SNOP). Other information about individual animals is also recorded -- such information as where they came from, which cages they were kept in, when they died (e.g. if they died naturally, or were sacrificed at the end of the experiment, or sacrificed earlier because they clearly would not survive), and so forth.

The outcome is that for each animal, we have a list of the lesions affecting them when they died. An example of a condensed listing of just

2025545858

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

the cancer-related lesions is appended. From such listings, we can perform various analyses and statistical tests to see whether the rate of cancer was increased at any site or for any type of cancer.

The simplest sort of analysis can be performed if all the animals survived for the whole length of the experiment -- and in practice the same sort of analysis is performed provided a reasonable fraction survived that long and provided there were not too many early deaths. In that case, we can simply list the dose groups and the numbers of animals with tumors compared with the total number of animals examined;

Dose	Number with Tumor	Number Examined	
0 (control)	10	50	
0.5 MTD	25	50	for example
MTD	30	50	

However, things are not this simple. Similar results are available for

- o Many different sites (See below)
- o Many different tumor types (for examples)
- o Combinations of these

To determine whether the rate of cancer has been increased involves comparing the proportion with tumor in the control group with the proportion with tumor in the dosed groups, and deciding whether there is a significant increase in any dosed group(s). The choice of which sites and/or types of tumors to combine before performing such statistical tests can be difficult. Generally, various grades of tumors (nodules, adenomas, carcinomas) may be combined for any given site. Table 2 gives an example of the sort of combination and testing which is performed.

In addition to the simple numbers of animals with tumor, there is additional information available which may be used in more complicated cases. The date of death of each animal is recorded, and may be taken into account in time-adjusted analyses of tumor incidence and in the life-table tests mentioned on the appended material.

For risk assessment purposes, it is necessary to make various assumptions about the behavior of animals in experiments like these. For example, it is assumed that:

- o Animals are affected independently (a tumor in one animal has no effect on any other animal).
- o Animals are equally likely to be affected
- o Each animal receives the same dose

.....

It is assumed that cage effects, littermate effects, the effects of heating, lighting, stress etc. are either not present, or are randomized among all the animals in such a way that there will be no effect on the final analysis.

With such assumptions, the probability of an animal having a tumor is related to the dose by some sort of dose-response relationship, so that at any given dose this probability can be computed. The observed results, a number of animals with tumor out of a larger number examined,

20255455859

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

is then a binomial sample with this probability. In practice, we don't know what the dose-response relationship is - we wish to estimate it from the results. But we assume that we know the SHAPE of the dose-response relationship (specified by a mathematical formula), so that all that is required is to estimate some PARAMETERS in the mathematical formula.

For example, the E.P.A. uses a dose-response relationship of the form:

$$p = 1 - \exp\{ - (q_0 + q_1.d + q_2.d^2 + ... + q_{k-1}.d^{k-1})\}$$

when there are k doses in an experiment, where p is the lifetime probability of tumor at dose d. It is usual to use a maximum likelihood technique to estimate the various parameters q_0, q_1, q_2, ... q_{k-1}, given the observed numbers of animals with tumors and the numbers of animals examined at each dose.

In cases where there is appreciable early mortality in the experiment, so that the observed numbers of animals with tumors are likely to be underestimates of what would have been observed at the end of a perfect experiment, one can make modifications to the dose response relationship, just as one can make life-table adjustments to standard statistical tests. One technique used is to modify the dose response curve to explicitly include length of life, using the idea that probability of tumor is likely to increase with a power of age (see page 1):

$$p = 1 - \exp\{ - (q_0 + q_1.d + q_2.d^2 + ... + q_{k-1}.d^{k-1}) \ (t/L)^n\}$$

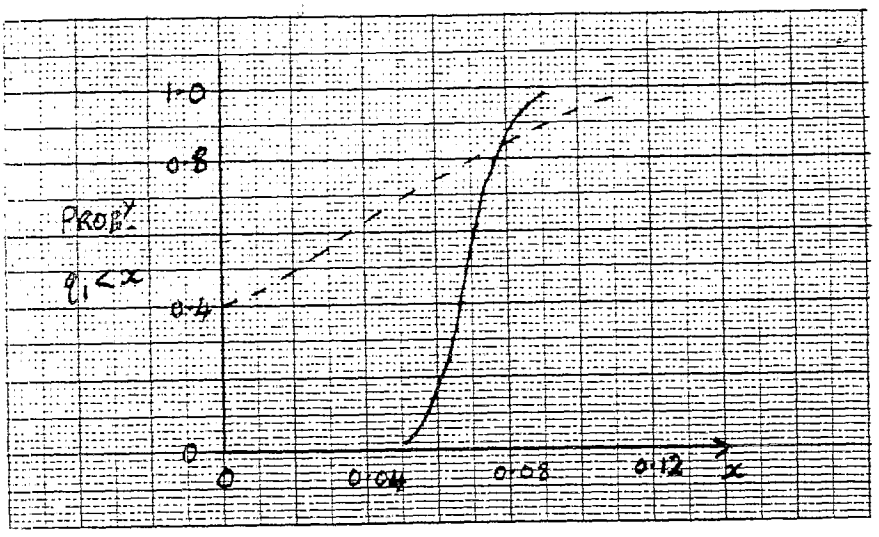
where t is the age at death, and L is a standard lifetime. The parameter n can either be fixed at some reasonable value (in the range 2 to 11), or estimated from the experimental results. This technique suffers from the same limitations as the usual modifications to the standard statistical tests -- one has to introduce additional assumptions in order to apply it. In this case, one has to decide whether the tumors were a cause of death, or simply incidental.

An alternative technique used when there is early mortality is to estimate the age dependence directly from the data, using a (so-called) non-parametric technique. This approach has been used to assemble a large database of comparable analyses of animal bioassays.

This methodology has taken the raw results of the animal experiment, and summarized them in the form of a dose-response curve with known parameters. It is also possible to estimate how uncertain one is about a given parameter, using the same maximum likelihood techniques used to obtain point estimates of them - indeed, one can plot the uncertainty distribution for any of the parameters. For example, for the parameter q_1 (which will turn out to be the one of interest), we can plot the probability that q_1 lies below any given value:

2025545860

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

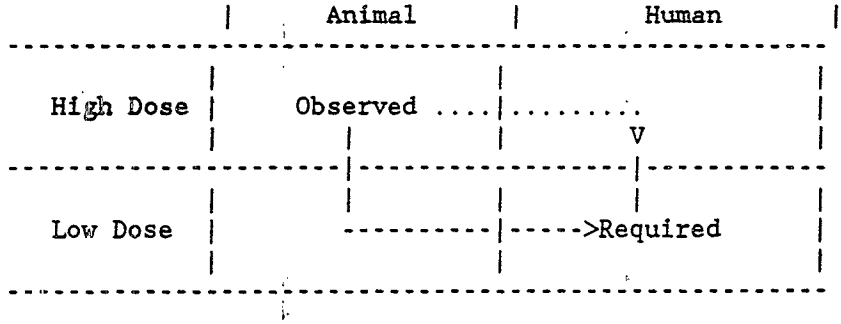


In particular, we can find that value q_1^* such that there is 95% probability that $q_1 < q_1^*$.

However, it is important to note that the uncertainty distribution so plotted contains only the uncertainty due to the numerical size of the experiment -- the uncertainty that arises because we used a small number of animals, instead of an infinite number. It does not include the uncertainties which must be present because of the shakiness of all our assumptions.

7. The Two Major Extrapolations

The assumptions made so far have allowed us to parametrize an animal dose-response relationship, obtaining values for the parameters which are presumably reasonably appropriate for high doses. Strictly speaking, this parametrization of the dose-response curve only enables us to estimate the results we would expect to see at high doses in animals - the dose-resposne relationship can only be relied on to interpolate between high doses and perhaps to extrapolate a short distance outside the experimental range of doses. The problem now is to perform two extrapolations - from animals to humans, and from high dose to low dose:



LOGICALLY there are two distinct routes to follow in this extrapolation, since there are logically two distinct dose-response

2025545861

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

curves involved (see below). One can extrapolate from high dose to low dose using the ANIMAL dose-response curve, and then extrapolate to humans (dashed lines), or extrapolate to humans at high doses and then use a HUMAN dose-response curve to extrapolate to low doses.

We have seen how to estimate the parameters of the (high dose region of) the animal dose-response curve. In practice, the same curve (with the same parameters) is is used to extrapolate to low doses, by building into the mathematical structure of the dose-response curve all our assumptions about low dose behavior.

How is this relevant for estimating human risk? Consider a generalized situation in which we wish to estimate the response (R) of humans to some dose (D) of material, when there is a response (r) in some animal at dose (d). Notice that nothing implies that r, R measure the same sort of response - they could be completely different (r could be acute toxicity to the lung of a mouse, R could be skin rashes in humans). Similarly, the dose measures d, D may be completely different. In the case immediately at hand, r is the lifetime probability of tumor in animals, and d is a dose as measured in the animal experiment. There are other cases of practical importance however - r might be some measure of response (such as number of revertants per culture dish) in a mutagenesis bioassay, with d the dose applied to each culture dish.

	Animal	Human
Response:	r	R
	(lifetime probability of tumor, p)	
Dose measure:	d	D
	(as used in experiments)	
Dose-response curve:		
	$r = f(d;\ a,b,c,\ldots t)$	$R = F(D;\ A,B,C,\ldots .T)$
	$[p = 1 - \exp\{-(q_0+q_1.d+\ldots)\}\ ]$	

What is required is some connection between the parameters a,b,c,... of the animal dose-response relationship and the parameters A,B,C,... of the human dose-response relationship. These parameters presumably include those mentioned in section 6, and I have explicitly included age amongst them. Given such a connection, the extrapolation to humans of the results in the animal studies is perfectly straightforward. The problem lies in finding the connection.

Once such a connection is found (by whatever means) we have the methodology for the two extrapolations required. Notice the difference between what is done in the two distinct pathways of extrapolation mentioned above:

In the first, the shape of the dose-response curves are examined, and it is decided how they may be (separately) extrapolated to low doses. Then some relationship is postulated between the parameters of the dose-response curves at low doses (it has to be postulated, since nothing can be measured at such low doses). One potential advantage of this approach is that the animal dose-response curve could be measured, in principle and by heroic experimentation, down to lower response rates than usual (and this has been done in some cases) - allowing greater confidence in this extrapolation to low dose.

2025545862

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

In the second, some relation between the parameters of the dose-response curves is obtained at high doses (and this may be done experimentally, in principle, since at high doses the responses are measurable). Then it is decided how the human dose-response curve should be extrapolated to low doses. The advantage here is the possibility of direct comparison between species, albeit at high dose.

The difference between these two logically distinct routes of extrapolation might be important in some circumstances. For cancer risk assessment based on animal carcinogenesis bioassays, however, the distinction is glossed over (one might even say, ignored), by the practice of assuming the same mathematical form for the dose-response curve in both humans and animals (or more generally, in all species), and interpreting the parameters in the same way for both compared species.

In the general case, however, what is required is some sort of relationship between the parameters of the dose-response curves:

Animal	Human
$r = f(d;\ a,b,c...t)$	$R = F(D;\ A,B,C...T)$

We need to be able to derive the parameters A,B,C... from the values a,b,c which can be estimated from experiments, and then use the human dose-response curve to extrapolate to low doses.

The practical approach is to seek parametrizations of the dose-response curve which result in the derivation of A,B,C... being simple given a,b,c... Consider the case of acute toxicity, for example. It is found that the shape of the dose-response curve for acute toxicity, in which the response is death, is very similar for a large number of toxins and for many different species. There is, in this case, a threshold-type dose-response curve which can be nicely parametrized by two values: the dose at which 50% of the animals tested can be expected to die (under suitable conditions), and the slope of the dose-response curve at this dose. The first parameter is known as the LD_{50} (the second has no special name).

Why is this parametrization useful? If the LD_{50}s of various materials in one species are plotted against the LD_{50}s of the same materials in another species, one finds approximate proportionality between them (the plot is a straight line). This can be expressed as, for example,

$$LD_{50}(\text{rabbit}) \text{ is proportial to } LD_{50}(\text{mouse}).$$

Even more remarkable, it turns out that if the dose is measured in a suitable way, as (amount)/(surface area of animal), then approximately we actually have

$$LD_{50}(\text{rabbit}) = LD_{50}(\text{mouse}) = LD_{50}(\text{other species})$$

and it is this approximate equality which explains the utility of the LD_{50}. The other parameter used in defining the dose-response curve, the slope of the curve at the LD_{50}, is not involved in this relationship.

20255845863

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

Had we chosen some other method of parametrization, it is quite possible the required interspecies relationship between parameters would be much more complicated.

8. Interspecies Comparison - Constant Relative Potency

What is sought is a simple relationship between the parameters of dose-response relationships in different species. When it is assumed that the dose-response relationship includes a term linear in dose, there is a simple measure of the strength of a carcinogen - the carcinogenic potency (the slope of the dose-response curve at low dose). The simplest hypothesis is that for different species, the ratio of carcinogenic potencies is constant for different materials, so that if material A is twice as potent a carcinogen as material B in species 1, it will also be twice as potent as material A in species 2. This is the idea of constant relative potency, as applied to carcinogenesis, and it underlies the standard approaches to estimating human risks from animals.

There is even some data which supports this idea! There have been several hundred bioassays performed simultaneously on rats and mice, and when the results of these are parametrized using a dose-response relationship which includes a linear term, we can estimate the potency in two species for each material tested. Plotting the potency measured in rats versus the potency measured in mice for each material then gives the figure shown. Notice that each measurement is uncertain to greater or lesser degree, due to the relatively small numbers of animals tested. If the idea of constant relative potency were exactly correct, these points would all lie on a straight line on the figure - or at least, all would lie sufficiently close to such a line that the measurement uncertainty bars on each point would encompass the line. From the figures, one can see that:

(1) On average, potency in one species is proportional to potency in the other species.

(2) There is a large scatter of the points around the lines of exact proportionality - a scatter bigger than would be expected from the measurement errors alone.

A similar comparison can be attempted between the potencies measured in animal experiments, and those observed in humans. These cases have arisen in the past where humans have been exposed to materials before they were known to be carcinogenic. We can make use of other's misfortune to estimate how potent each such material is in humans, and compare with estimates obtained for mice and rats in laboratory experiments. In this case, the uncertainties are so large that little can be quantitatively states, although qualitatively the idea of constant relative potency does not seem to be disproved. A more recent and much more thorough study of comparisons between humans and animals has been carried out for the E.P.A. by Dr. Kenny Crump, and we can expect that to be published soon - I understand that conclusions are qualitatively similar.

2025545864

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

9. Interspecies comparisons - practical and theoretical

The measure of carcinogenic potency introduced above was roughly defined as the ratio of (excess tumor probability)/(dose), at low enough dose. For the E.P.A. model usually used in risk assessments:

$$p = 1 - \exp\{ - (q_0 + q_1.d + q_2.d^2 + \ldots + q_{k-1}.d^{k-1})\}$$

the corresponding measure is q_1. When this dose-response relationship is used with real data, it is usual to use an "upper 95% confidence limit" estimate q_1^* of q_1 as the measure of potency, since such an estimate is always non-zero (while, for example, the maximum likelihood estimate is often zero). The "upper 95% confidence limit" is with respect to the numerical uncertainties of the experiment only, and so this estimate of potency is in no sense an upper limit with respect to all the other uncertainties involved.

To compare humans with animals, the approach taken is to use a similar dose-response relationship in both cases:

Animal	Human
$p = 1 - \exp\{ - (q_0 + q_1.d + \ldots)\}$	$p = 1 - \exp\{ - (Q_0 + Q_1.D + ..)\}$

and then the constant relative potency hypothesis suggests that Q_1 is proportional to q_1, or to our estimate q_1^* of it:

$$Q_1 = \text{const. } q_1^*$$

where the constant depends only on which animals species is used. We expect the constant to be different for different animal species - it will presumably depend on how we measure dose, on the relative lifespans of animal and human, on relative metabolic rates, and a whole host of other factors. With enough experiments, we could measure the constant in this relationship - at least in comparing animal with animal, rather than human with animal - and (in theory) empirically determine how it varies with these factors. The graphs above suggest that the constant is not completely constant, but that there is some sort of random uncertainty built in (or at least, an uncertainty that we can treat as random), amounting to an average factor of about 5.

If we are very lucky, it may be possible to find some way of measuring dose so that the constant in the above relationship is numerically equal to 1, so that the potency is equal in different species (up to the uncertainties) - just as it was possible to find such a measure in the case of the LD_{50}.

In practice, the E.P.A. assumes that the constant is exactly unity if the dose is measured as a (daily average amount)/(surface area of animal), by analogy with the LD_{50} case. (The graphs shown above actually suggest that it would be better to assume an average factor of unity, with an uncertainty factor of about 5, when the dose is measured as a (daily average amout)/(bodyweight of animal)).

2025545865

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

10. An example - 1,2 Dibromoethane

As an example of the procedures usually adopted, let us look at the case of 1,2-Dibromoethane. What follows is by now means complete, but it indicates the sort of analysis which has to be performed. This example is confined to analysing just one result out of many, in a single bioassay (of about 5). In practice, it is essential to look at all the results.

The bioassay I have chosen was an inhalation bioassay in the National Toxicology Program series. A summary of the study design is:

	Initial number of animals	Concentration ppm (6 hrs/d, 5 d/wk)	Time on study exposed (weeks)	Time on study observed (weeks)
Male rats				
control	50	0	0	104-106
low-dose	50	10	103	1
high-dose	50	40	88	0-1
Female rats				
control	50	0	0	104-106
low-dose	50	10	103	1
high-dose	50	40	91	0-1

And similarly for mice

We will look only at the results in female rats. First, their survival was not as good as might be desired in such an experiment, but the early mortality was probably due to the cancers appearing in the study, so it is acceptable - we can use (at least initially) the simplest analysis based on "end-of-life" data, without having to worry too much about the age dependence (this should always be backed up by further analysis, of course).

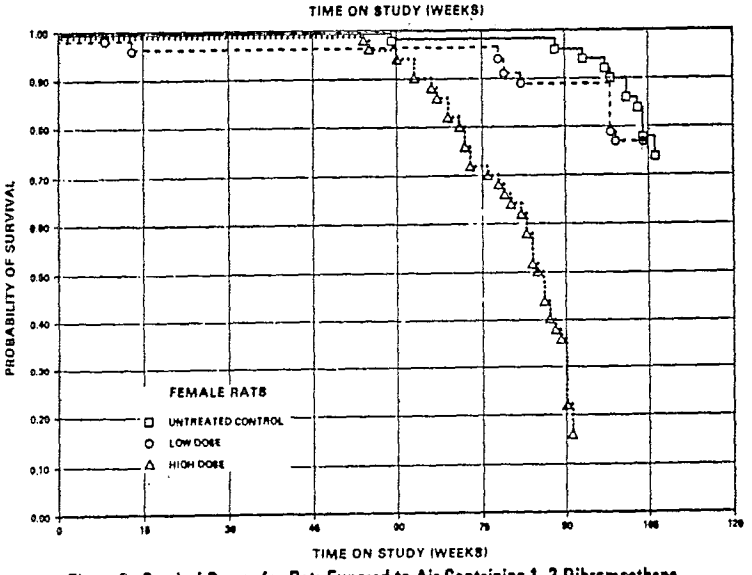


Figure 2. Survival Curves for Rats Exposed to Air Containing 1, 2-Dibromoethane

2025545866

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

Tumors were found in many tissues. A summary of those tissues where more than 5% of the animals in any group were found with tumors is (for female rats):

	Control	Low	High
Subcutaneous tissue: fibroma	0/50	0/50	3/50
Subcutaneous tissue: fibroma or fibrosarcoma	0/50	0/50	4/50
Nasal Cavity: Carcinoma, NOS	0/50	0/50	25/50
Nasal Cavity: Sqamous cell carcinoma	1/50	1/50	5/50
Nasal Caavity: Adenoma, NOS	0/50	11/50	3/50
Nasal Cavity: Adenocarcinoma, NOS	0/50	20/50	29/50
Nasal Cavity: Adenomatous Polyp, NOS	0/50	5/50	5/50
Nasal Cavity: Papillary Adenoma	0/50	3/50	0/50
Nasal Cavity: Adenoma,NOS; Carcinoma, NOS; Adenocarcinoma,NOS; Papillary Adenoma; Adenomatous polyp,NOS; and Sqamous cell Carcinoma	1/50	34/50	43/50
Lung: Alveolar/Bronchiolar Carcinoma	0/50	0/48	4/47
Lung: Alveolar/Bronchiolar Carcinoma or Adenoma	0/50	0/48	5/47
Hematopoietic System: All leukemias	6/50	7/50	1/50
Hematopoietic System: Monocytic leukemia	6/50	5/50	1/50
Circulatory System: Hemangiosarcoma	0/50	0/50	5/50
Circulatory System: Hemangiosarcoma or Hemangiosarcoma, invasive	0/50	0/50	5/50
Liver: Neoplastic nodule	2/50	0/49	3/48
Liver: Hepatocellular carcinoma	0/50	1/49	3/48
Liver: Neoplastic nodule or Hepatocellular carcinoma	2/50	1/49	5/48
Pituitary: Adenoma, NOS	1/50	18/49	4/45
Pituitary: Chromophobe adenoma	20/50	0/49	0/45
Adrenal: Pheochromocytoma	3/50	1/49	0/47
Thyroid: C-cell Carcinoma	1/49	3/48	1/45
Mammary Gland: Adenocarcinoma,NOS	1/50	0/50	4/50
Mammary Gland: Fibroadenoma	4/50	29/50	24/50

Notice especially the various groupings which are employed - this is a matter of judgement. It is clear that the major effect is in the nasal cavity, but observe also the effect on fibroadenomas in the mammary gland, and the **negative** trend seen in the pituitary. Such negative trends are **ignored**.

Using the combined results in the nasal cavity, we fit the E.P.A. multistage model and find best estimates of:

$$q_0 = 2.699 \times 10_{-2};\ q_1 = 6.876 \times 10^{-2};\ q_2 = 0;$$

and obtain an upper confidence limit for q_1 of $q_1^* = 8.6 \times 10_{-2}$ in all cases using as doses the values 0, 10 and 40 from the experimental design. In fact, the earlier figure of a distribution of values for q_1 is taken from this example - you can read the probability of q_1 being less than any given value from that figure.

20255458 67

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

Now what do we do with this estimate? That depends on the application, but we will assume that we wish to make a "UNIT RISK" estimate for humans from it - that is, estimate the lifetime risk to a human exposed to 1 microgram/m^3 of dibromoethane for life.

There are several extrapolations required. First, the animals were dosed for a lifetime, but not continuously. Correcting for continuous exposure introduces a factor of 7/5 x 24/6 (for days/week and hours/day) - but notice the subtle assumptions being made here, that it is average exposure that matters (and not peak exposure, for example).

Now we estimate that a female rat will suffer and increased lifetime risk of about 0.48 per ppm in the air (we assume that we are talking about such low doses that the excess risk is small). 1 ppm for 1,2-dibromoethane corresponds to about 7.6 mg/m^3 (one would estimate a little higher from the perfect gas laws), or 7600 microgram/m^3, so that the increased lifetime risk to a female rat exposed continuously to 1 migrogram/m^3 is about 6.3 x 10^{-5}

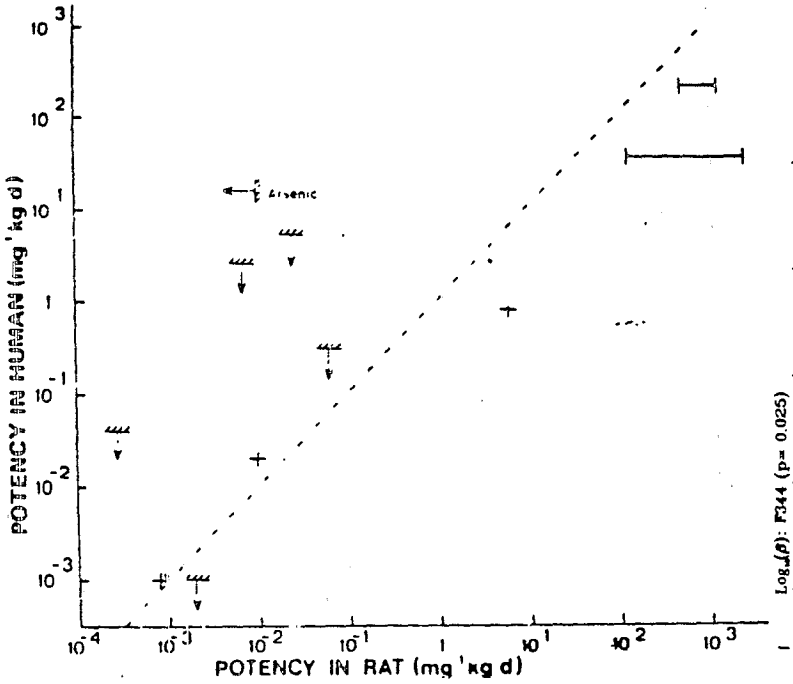
What about humans? We saw before that the assumption made was that humans are just as sensitive as animals - i.e. they suffer equal lifetime risks - if exposed at doses which are equal on an (amount)/(surface area) basis. Now it turns out that, approximately, equal concentrations in air lead to exposures which are equivalent on this basis, provided the species under consideration absorb about the same amount from the air they breathe. Thus the extrapolation to humans is simple in this case - one simply takes the same value for humans - a "UNIT RISK" of about 6.3 x 10^{-5} (i.e. that is the lifetime risk from continuous exposure to 1 migrogram/m^3 of dibromoethane in the air).

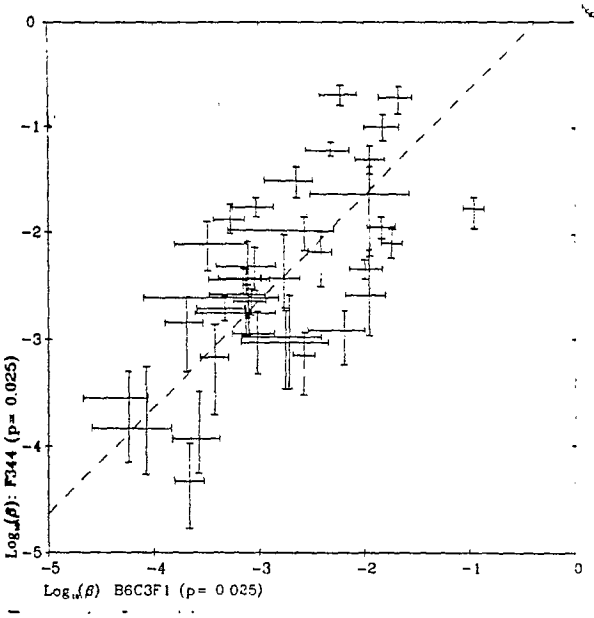
It may be desired to estimate from this the effect on humans of ingestion of dibromoethane. In this case there are actually other bioassays in which dibromoethane was fed to animals under various conditions, but suppose that we have to make some estimate from the inhalation data. The "standard" human inhales, on average, about 20 m^3 of air per day, and so inhales about 20 microgram/day of contaminant from air contaminated with 1 microgram/m^3. If we assume that 100% of this contaminant is absorbed, the human's daily dose is 20 micrograms/day, or about 20/70 microgram/kg-day (as a fraction of bodyweight), or 2.9 x 10^{-4} mg/kg-day in the conventional units used. This results in a risk of about 6.3 x 10^{-5}, as detailed above, so that the potency is just the ratio of these -- 0.22 $(mg/kg\text{-}day)^{-1}$.

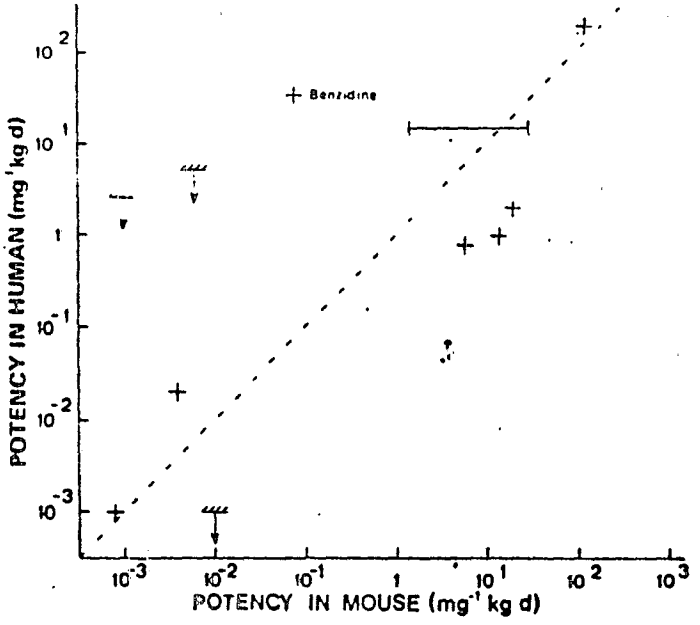
These short outline calculations have made several assumptions which require examination in any particular case. We have not looked at all the bioassay results, so one cannot expect that the numbers obtained here will correspond with what anybody else, who has done a more thorough job, will obtain -- they are placed here in order to show in outline what is done. In practice, one has to decide that the tumor site and type combinations are appropriate for combination in the animal species. That these tumors are relevant end points for estimating the probable effects on humans. That the route of administration, and method of administration are reasonable to produce results that may be extrapolated to humans. And a myriad of other details which have only been lightly touched upon, or completely omitted, in this sketch.

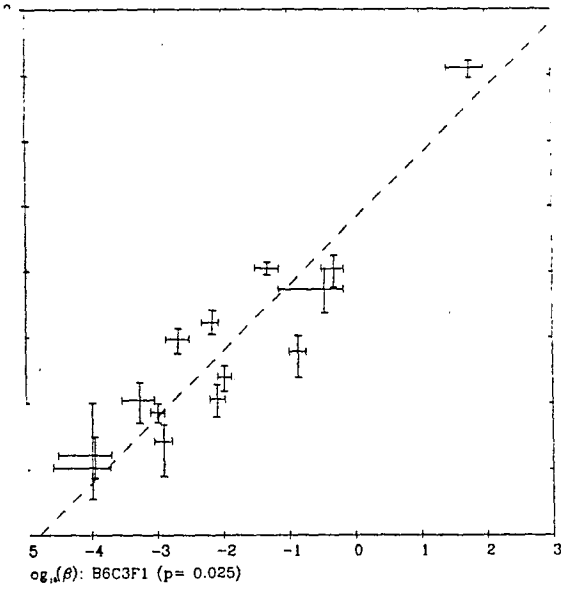
2025545868

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000









2025545869

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

TABLE E3. ANALYSIS OF PRIMARY TUMORS IN MALE MICE (Continued)

	Vehicle Control	500 mg/kg	1,000 mg/kg
Circulatory System: Hemangiosarcoma			
Overall Rates *(a)*	4/50 (8%)	3/49 (6%)	1/50 (2%)
Adjusted Rates *(b)*	10.1%	8.8%	2.6%
Terminal Rates *(c)*	3/38 (8%)	2/33 (6%)	1/39 (3%)
Life Table Tests *(d)*	P=0.130N	P=0.559N	P=0.169N
Incidental Tumor Tests *(d)*	P=0.097N	P=0.408N	P=0.176N
Cochran-Armitage Trend Test *(d)*	P=0.134N		
Fisher Exact Tests		P=0.512N	P=0.181N
Circulatory System: Hemangioma or Hemangiosarcoma			
Overall Rates *(a)*	4/50 (8%)	4/49 (8%)	1/50 (2%)
Adjusted Rates *(b)*	10.1%	11.8%	2.6%
Terminal Rates *(c)*	3/38 (8%)	3/33 (9%)	1/39 (3%)
Life Table Tests *(d)*	P=0.142N	P=0.579	P=0.169N
Incidental Tumor Tests *(d)*	P=0.110N	P=0.573N	P=0.176N
Cochran-Armitage Trend Test *(d)*	P=0.147N		
Fisher Exact Tests		P=0.631	P=0.181N
Liver: Adenoma			
Overall Rates *(a)*	0/50 (0%)	5/49 (10%)	13/50 (26%)
Adjusted Rates *(b)*	0.0%	13.0%	33.3%
Terminal Rates *(c)*	0/38 (0%)	3/33 (9%)	13/39 (33%)
Life Table Tests *(d)*	P<0.001	P=0.030	P<0.001
Incidental Tumor Tests *(d)*	P<0.001	P=0.023	P<0.001
Cochran-Armitage Trend Test *(d)*	P<0.001		
Fisher Exact Tests		P=0.027	P<0.001
Liver: Carcinoma			
Overall Rates *(a)*	10/50 (20%)	14/49 (29%)	12/50 (24%)
Adjusted Rates *(b)*	24.3%	35.9%	25.8%
Terminal Rates *(c)*	7/38 (18%)	9/33 (27%)	5/39 (13%)
Life Table Tests *(d)*	P=0.427	P=0.183	P=0.463
Incidental Tumor Tests *(d)*	P=0.536	P=0.379	P=0.548N
Cochran-Armitage Trend Test *(d)*	P=0.363		
Fisher Exact Tests		P=0.224	P=0.405
Liver: Adenoma or Carcinoma			
Overall Rates *(a)*	10/50 (20%)	18/49 (37%)	23/50 (46%)
Adjusted Rates *(b)*	24.3%	45.1%	49.8%
Terminal Rates *(c)*	7/38 (18%)	12/33 (36%)	16/39 (41%)
Life Table Tests *(d)*	P=0.013	P=0.042	P=0.014
Incidental Tumor Tests *(d)*	P=0.009	P=0.098	P=0.019
Cochran-Armitage Trend Test *(d)*	P=0.004		
Fisher Exact Tests		P=0.052	P=0.005
Forestomach: Squamous Cell Papilloma			
Overall Rates *(a)*	3/49 (6%)	3/48 (6%)	9/49 (18%)
Adjusted Rates *(b)*	7.9%	9.1%	23.1%
Terminal Rates *(c)*	3/38 (8%)	3/33 (9%)	9/39 (23%)
Life Table Tests *(d)*	P=0.038	P=0.597	P=0.065
Incidental Tumor Tests *(d)*	P=0.038	P=0.597	P=0.065
Cochran-Armitage Trend Test *(d)*	P=0.034		
Fisher Exact Tests		P=0.651	P=0.060

2025545870

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000

TABLE B3.

INDIVIDUAL ANIMAL TUMOR PATHOLOGY OF MALE MICE IN THE 2-YEAR STUDY OF BENZYL ACETATE

HIGH DOSE

ANIMAL NUMBER / WEEKS ON STUDY	TOTAL TISSUES / TUMORS
RESPIRATORY SYSTEM	
LUNGS AND BRONCHI	50
HEPATOCELLULAR CARCINOMA, METAS	1
ALVEOLAR/BRONCHIOLAR ADENOMA	6
ALVEOLAR/BRONCHIOLAR CARCINOMA	2
TRACHEA	49
HEMATOPOIETIC SYSTEM	
BONE MARROW	50
SPLEEN	49
HEMANGIOSARCOMA	1
LYMPH NODES	50
MALIGNANT LYMPHOMA, MIXED TYPE	1
THYMUS	49
CIRCULATORY SYSTEM	
HEART	50
DIGESTIVE SYSTEM	
SALIVARY GLAND	49
LIVER	50
NEOPLASM, NOS	1
HEPATOCELLULAR ADENOMA	13
HEPATOCELLULAR CARCINOMA	12
BILE DUCT	50
GALLBLADDER & COMMON BILE DUCT	50M
PANCREAS	49
ESOPHAGUS	48
STOMACH	49
SQUAMOUS CELL PAPILLOMA	9
SQUAMOUS CELL CARCINOMA	2
SMALL INTESTINE	47
LARGE INTESTINE	46
URINARY SYSTEM	
KIDNEY	50
TUBULAR-CELL ADENOMA	1
TUBULAR-CELL ADENOCARCINOMA	1
URINARY BLADDER	49
ENDOCRINE SYSTEM	
PITUITARY	46
ADRENAL	49
GANGLIONEUROMA	1
THYROID	47
FOLLICULAR-CELL ADENOMA	4
PARATHYROID	44
PANCREATIC ISLETS	49
ISLET-CELL ADENOMA	2
REPRODUCTIVE SYSTEM	
MAMMARY GLAND	50M
TESTIS	50
INTERSTITIAL-CELL TUMOR	2
PROSTATE	49
NERVOUS SYSTEM	
BRAIN	50
SPECIAL SENSE ORGANS	
HARDERIAN GLAND	50M
ADENOMA, NOS	3
BODY CAVITIES	
MESENTERY	50M
HEPATOCELLULAR CARCINOMA, METAS	1
ALL OTHER SYSTEMS	
MULTIPLE ORGANS NOS	50M
HEPATOCELLULAR CARCINOMA, METAS	1
MALIGNANT LYMPHOMA, NOS	1
MALIG.LYMPHOMA, LYMPHOCYTIC TYP	1

M ANIMALS NECROPSIED

+: TISSUE EXAMINED MICROSCOPICALLY
-: REQUIRED TISSUE NOT EXAMINED MICROSCOPICALLY
X: TUMOR INCIDENCE
M: NECROPSY, NO AUTOLYSIS, NO MICROSCOPIC EXAMINATION
S: ANIMAL MIS-SEXED
.: NO TISSUE INFORMATION SUBMITTED
C: NECROPSY, NO HISTOLOGY DUE TO PROTOCOL
A: AUTOLYSIS
M: ANIMAL MISSING
B: NO NECROPSY PERFORMED

ETC.

202554587

Source: https://www.industrydocuments.ucsf.edu/docs/kzbl0000